

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/594,064
Source: IFWIP
Date Processed by STIC: 10/3/06

ENTERED



IFWP

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/594,064

DATE: 10/03/2006

TIME: 08:43:02

Input Set : A:\21571P SEQLIST.TXT
 Output Set: N:\CRF4\10032006\J594064.raw

```

4 <110> APPLICANT: Bryan, Janine T.
5   Brownlow, Michelle K.
6   Schultz, Loren D.
7   Janse, Kathrin U.
9 <120> TITLE OF INVENTION: OPTIMIZED EXPRESSION OF HPV 52 L1 IN
10  YEAST
12 <130> FILE REFERENCE: 21571P
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/594,064
C--> 14 <141> CURRENT FILING DATE: 2006-09-25
14 <150> PRIOR APPLICATION NUMBER: PCT/US2005/009199
15 <151> PRIOR FILING DATE: 2005-03-18
17 <150> PRIOR APPLICATION NUMBER: 60/555926
18 <151> PRIOR FILING DATE: 2004-03-24
20 <160> NUMBER OF SEQ ID NOS: 7
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 1512
26 <212> TYPE: DNA
27 <213> ORGANISM: Artificial Sequence
29 <220> FEATURE:
30 <223> OTHER INFORMATION: HPV52L1R
32 <400> SEQUENCE: 1
33 atgtccgtct ggagaccatc cgaagctact gtctacttgc caccagttcc agtctctaag 60
34 gttgtctcta ccgacgaata cgtctccaga acctccatct actactacgc tggttcctct 120
35 agattgtga ctgtcggtca cccatacttc tctatcaaga acaccccttc cggttaacgg 180
36 aagaaggctt tggttccaaa ggtctctggt ttgcaataca gagtcttcag aatcaagttg 240
37 ccagacccaa acaagttcgg tttcccagac actagttct acaacccaga aactcaaaga 300
38 ttggtgtctggg cttgtactgg ttggaaatc ggttagaggc aaccattggg tgtcggtatc 360
39 tctggtcacc cattgttgaa caagttcgac gacactgaaa cctctaacaa gtacgctgg 420
40 aaggccaggta tcgataaacag agaatgtttg tctatggact acaagcaaac tcaattgtgt 480
41 atcttgggtt gtaagccacc aatcggtgaa cactgggttta agggtactcc atgtaacaac 540
42 aactctggta acccaggtga ctgtccacca ttgcaattga tcaactccgt catccaagac 600
43 ggtgacatgg tcgacactgg ttcgggttgt atggacttca acacccgtca agcttctaag 660
44 tccgacgtcc caatcgacat ctgttccctt gtctgttaatg accccagacta cttgcaaatg 720
45 gcttctgaac catacggtga ctccctgttc ttcttcttga gaagagaaca aatgttcgtc 780
46 agacacttct tcaacagagc tggcaccttgg ggtgacccag ttccagggtga cttgtacatc 840
47 caagggttcca actctgttaa cactgtact gtcacatccct ctgtttctt cccaaactcca 900
48 tctgggttcca tggtcacccat cgaatccccaa ttgttcaaca agccatactg gttgaaaga 960
49 gctcaagggtc acaacaacgg tatctgttg ggttaaccat tggacttgc acgtcgac 1020
50 actactagat ctactaacat gacccgttgt gctgaagtca agaaggaaatc cacccataac 1080
51 aacgaaaaact tcaaggaataa cttgagacac ggtgaagaat tggacttgc attcatcttc 1140
52 caattgtgtta agatcacctt gaccgctgac gtcatgactt acatccacaa gatggacgct 1200
53 actatcttgg aagactggca attcggttttgc actccaccac catccgcttc cttggaaagac 1260

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/594,064

DATE: 10/03/2006
TIME: 08:43:02

Input Set : A:\21571P SEQLIST.TXT
Output Set: N:\CRF4\10032006\J594064.raw

54 acttacagat tcgtcacttc cactgctatc acctgtcaaa agaacactcc accaaagggt 1320
55 aaggaagacc cattgaagga ctacatgttc tgggaagtgc acttgaagga aaagttctct 1380
56 gctgacttgg accaatccc attgggtaga aagttcttgt tgcaagctgg tttgcaagct 1440
57 agacccaaagt tgaagagacc agctagctct gctccaagaa cttccaccaa gaagaagaag 1500
58 gtcaagagat aa 1512
60 <210> SEQ ID NO: 2
61 <211> LENGTH: 503
62 <212> TYPE: PRT
63 <213> ORGANISM: Human papillomavirus type 52
65 <400> SEQUENCE: 2
66 Met Ser Val Trp Arg Pro Ser Glu Ala Thr Val Tyr Leu Pro Pro Val
67 1 5 10 15
68 Pro Gly Ser Lys Val Val Ser Thr Asp Glu Tyr Val Ser Arg Thr Ser
69 20 25 30
70 Ile Tyr Tyr Tyr Ala Gly Ser Ser Arg Leu Leu Thr Val Gly His Pro
71 35 40 45
72 Tyr Phe Ser Ile Lys Asn Thr Ser Ser Gly Asn Gly Lys Lys Val Leu
73 50 55 60
74 Val Pro Lys Val Ser Gly Leu Gln Tyr Arg Val Phe Arg Ile Lys Leu
75 65 70 75 80
76 Pro Asp Pro Asn Lys Phe Gly Phe Pro Asp Thr Ser Phe Tyr Asn Pro
77 85 90 95
78 Glu Thr Gln Arg Leu Val Trp Ala Cys Thr Gly Leu Glu Ile Gly Arg
79 100 105 110
80 Gly Gln Pro Leu Gly Val Gly Ile Ser Gly His Pro Leu Leu Asn Lys
81 115 120 125
82 Phe Asp Asp Thr Glu Thr Ser Asn Lys Tyr Ala Gly Lys Pro Gly Ile
83 130 135 140
84 Asp Asn Arg Glu Cys Leu Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys
85 145 150 155 160
86 Ile Leu Gly Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Thr
87 165 170 175
88 Pro Cys Asn Asn Ser Gly Asn Pro Gly Asp Cys Pro Pro Leu Gln
89 180 185 190
90 Leu Ile Asn Ser Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe
91 195 200 205
92 Gly Cys Met Asp Phe Asn Thr Leu Gln Ala Ser Lys Ser Asp Val Pro
93 210 215 220
94 Ile Asp Ile Cys Ser Ser Val Cys Lys Tyr Pro Asp Tyr Leu Gln Met
95 225 230 235 240
96 Ala Ser Glu Pro Tyr Gly Asp Ser Leu Phe Phe Leu Arg Arg Glu
97 245 250 255
98 Gln Met Phe Val Arg His Phe Phe Asn Arg Ala Gly Thr Leu Gly Asp
99 260 265 270
100 Pro Val Pro Gly Asp Leu Tyr Ile Gln Gly Ser Asn Ser Gly Asn Thr
101 275 280 285
102 Ala Thr Val Gln Ser Ser Ala Phe Phe Pro Thr Pro Ser Gly Ser Met
103 290 295 300
104 Val Thr Ser Glu Ser Gln Leu Phe Asn Lys Pro Tyr Trp Leu Gln Arg

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/594,064

DATE: 10/03/2006
TIME: 08:43:02

Input Set : A:\21571P SEQLIST.TXT
Output Set: N:\CRF4\10032006\J594064.raw

105	305	310	315	320
106	Ala Gln Gly His Asn Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val			
107		325	330	335
108	Thr Val Val Asp Thr Thr Arg Ser Thr Asn Met Thr Leu Cys Ala Glu			
109		340	345	350
110	Val Lys Lys Glu Ser Thr Tyr Lys Asn Glu Asn Phe Lys Glu Tyr Leu			
111		355	360	365
112	Arg His Gly Glu Glu Phe Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys			
113		370	375	380
114	Ile Thr Leu Thr Ala Asp Val Met Thr Tyr Ile His Lys Met Asp Ala			
115		385	390	395
116	Thr Ile Leu Glu Asp Trp Gln Phe Gly Leu Thr Pro Pro Pro Ser Ala			
117		405	410	415
118	Ser Leu Glu Asp Thr Tyr Arg Phe Val Thr Ser Thr Ala Ile Thr Cys			
119		420	425	430
120	Gln Lys Asn Thr Pro Pro Lys Gly Lys Glu Asp Pro Leu Lys Asp Tyr			
121		435	440	445
122	Met Phe Trp Glu Val Asp Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp			
123		450	455	460
124	Gln Phe Pro Leu Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Gln Ala			
125		465	470	475
126	Arg Pro Lys Leu Lys Arg Pro Ala Ser Ser Ala Pro Arg Thr Ser Thr			
127		485	490	495
128	Lys Lys Lys Lys Val Lys Arg			
129		500		
132	<210> SEQ ID NO: 3			
133	<211> LENGTH: 1512			
134	<212> TYPE: DNA			
135	<213> ORGANISM: Human Papillomavirus Type 52			
137	<400> SEQUENCE: 3			
138	atgtccgtgt ggcggcctag tgaggccact gtgtacctgc ctcctgtacc tgtctctaag 60			
139	gtttaagca ctgatgagta tggatctcgca acaagcatct attattatgc aggcagttct 120			
140	cgttactaa cagtaggaca tccctatttt tctattaaaa acaccatgg tggtaatgg 180			
141	aaaaaaagtt tagttccaa ggtgtctggc ctgcaataca gggattttttt aattaaattt 240			
142	ccggacccta ataaatttgg ttttccagat acatctttt ataaccagg aacccaaagg 300			
143	tgggtgtggg cctgtacagg cttggaaatt ggttagggac agcctttagg tgggttatt 360			
144	atggggcattt ctttattaaa caagttttagt gatactggaa ccagtaacaa atatgctgg 420			
145	aaacctggta tagataatag ggaatgtta tctatggatt ataaggcagac tcagttatgc 480			
146	attttaggat gcaaacctcc tataggtgaa cattggggta agggAACCC ttgtataat 540			
147	aatttcaggaa atcctgggaa ttgtcctccc ctacagctca ttaacagtgt aatacaggat 600			
148	ggggacatgg tagatacagg atttgggtgc atggattttt ataccttgc agctagtaaa 660			
149	atgtatgtgc ccattgatat atgttagcagt gtatgttgtt atccagatta tttgcaaatg 720			
150	gctagcgagc catatggta cagttgttc tttttctta gacgtgagca aatgtttgtt 780			
151	agacactttt ttaataggcc cggtagctta ggtgaccctg tgccaggtga ttatataata 840			
152	caagggtcta actctggcaa tactgccact gtacaaagca gtgtttttt tcctactcct 900			
153	atgggttcta tggtaacctc agaattccaa ttatttaata aaccgtactg gttacaacgt 960			
154	gcmcaggccacaataatgg catatgtgg ggcaatcagt tgggtgtcac agttgtggat 1020			
155	accactcgta gcactaacat gactttatgt gctgaggtt aaaaggaaag cacatataaaa 1080			
156	aatgaaaatt ttaaggaata cttcgatcat ggcgaggaat ttgattaca atttattttt 1140			

RAW SEQUENCE LISTING DATE: 10/03/2006
PATENT APPLICATION: US/10/594,064 **TIME:** 08:43:02

Input Set : A:\21571P SEQLIST.TXT
Output Set: N:\CRF4\10032006\J594064.raw

```

157 caatttgtca aaattacatt aacagctgat gttatgacat acattcataa gatggatgcc 1200
158 actatTTtag aggactggca atttggccctt acccccaccac cgtctgcata tttggaggac 1260
159 acatacagat ttgtcaCTTC tactgctata acttgcTaaa aaaacacGCC acctaaaggA 1320
160 aaggaagatC cttaaaAGGA ctatATgtt tgggaggTgg atttaaaAGA aaagtTTCT 1380
161 gcagatTTAG atcagTTCC tttaggtAGG aagtTTGT tacaggcagg gctacaggCT 1440
162 agGCCAAAC taaaACGCC tgcATCATCG GCCCACGTA CCTCCACAAA gaagaaaaAG 1500
163 gttaaaaggT aa 1512
165 <210> SEQ ID NO: 4
166 <211> LENGTH: 21
167 <212> TYPE: DNA
168 <213> ORGANISM: Artificial Sequence
170 <220> FEATURE:
171 <223> OTHER INFORMATION: PCR Primer
173 <400> SEQUENCE: 4
174 atgtccgtgt ggccgcctag t 21
176 <210> SEQ ID NO: 5
177 <211> LENGTH: 22
178 <212> TYPE: DNA
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: PCR Primer
184 <400> SEQUENCE: 5
185 gagatctcaa ttacacaaAG tg 22
187 <210> SEQ ID NO: 6
188 <211> LENGTH: 31
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: PCR Primer
195 <400> SEQUENCE: 6
196 gagatctcac aaaacaaaat gtccgtgtgg c 31
198 <210> SEQ ID NO: 7
199 <211> LENGTH: 1512
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: 52L1R antisense
206 <400> SEQUENCE: 7
207 tacaggcaga cctctggtag gttcgatga cagatgaacg gtggtcaagg tcagagattc 60
208 caacagagat ggctgcttat gcagaggTct tggaggtaga tcatgtatgcg accaaggaga 120
209 tctaacaact gacaggccagt gggtatgaag agatagtct tggaggAGG gccattGCC 180
210 ttcttccaga accaaggTT ccagagacca aacgttatgt ctcagaAGTC ttgttcaac 240
211 ggtctgggtt tggcaagcc aaagggtctg tggatcaaaga tgggggtct ttgagTTCT 300
212 aaccagacCC gaacatgacc aaacCTTtag ccatCTCCAG ttggtaACCC acGCCatAG 360
213 agaccagtgg gtaacaactt gttcaagCTG ctgtgacttt ggagatttt catgcgacca 420
214 ttccggTccat agtattgtc tcttacAAAC agataacCTGA tggccgttt agttAACACA 480
215 tagaACCCAA cattcggTgg ttggccactt gtgacCCcat tcccatgagg tacattgttG 540
216 ttgagaccat tgggtccact gacaggTgtt aacgttaact agttgaggca tgggttCTG 600
217 ccactgtacc agctgtgacc aaagCCAACA tacctgaagt tggaaacgt tcgaagattc 660

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/594,064

DATE: 10/03/2006
TIME: 08:43:02

Input Set : A:\21571P SEQLIST.TXT
Output Set: N:\CRF4\10032006\J594064.raw

218 aggctgcagg gttagctgta gacaaggaga cagacattca tgggtctgat gaacgtttac 720
219 cgaagacttg gtatgccact gaggaacaag aagaagaact ctctcttgt ttacaagcag 780
220 tctgtgaaga agttgtctcg accatggAAC ccactgggtc aaggTccact gaacatgtag 840
221 gttccaagggt tgagaccatt gtgacgatga caggttagga gacgaaagaa gggttggaggt 900
222 agaccaagggt accagtggag gcttagggTT aacaagtgt tcggatgac caacgtttct 960
223 cgagttccag tggTTTgCC atagacaacc ccattggta acaagcagt gcagcagctg 1020
224 tggatgatcta gatgattgta ctggAACACa cgacttcagt tcttccttag gtggatgtc 1080
225 ttgcTTTgA agttcTTtat gaaCTCTGT ccacttCTTA agCTGAACGT taagtAgAAAG 1140
226 gttAACACAT tctagtggaa ctggcgactg cagtaCTGAA tggtaggttt ctacCTGCGA 1200
227 tgatAGAACc ttctgaccgt taAGCCAAAC tgaggTggTg gtaggcgaag gAACCTTCTG 1260
228 tgaATgtcta agcagtgaag gtgacgatag tggacAGTT tcttgtaggg tggttcccA 1320
229 ttccTTCTGG gtaACTTCt gatgtacaAG accCTTCAGC tgaACTTCt tttcaAGAGA 1380
230 cgactgaacc tggTTAAGGG TAACCCATCT ttcaAGAACa acgTTcgacc aaACgttCGA 1440
231 tctggTTTCA acttCTCTGG tcgatcgaga cgaggttctt gaaggTggTT ctTCTTCTTC 1500
232 cagttctcta tt 1512

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/594,064

DATE: 10/03/2006

TIME: 08:43:03

Input Set : A:\21571P SEQLIST.TXT

Output Set: N:\CRF4\10032006\J594064.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date